J Clin Pathol 1987;**40**:210–215

Autoimmunity: a decision theory model

J A MORRIS

From the Department of Pathology, Lancaster Moor Hospital, Lancaster

SUMMARY Concepts from statistical decision theory were used to analyse the detection problem faced by the body's immune system in mounting immune responses to bacteria of the normal body flora. Given that these bacteria are potentially harmful, that there can be extensive cross reaction between bacterial antigens and host tissues, and that the decisions are made in uncertainty, there is a finite chance of error in immune response leading to autoimmune disease. A model of ageing in the immune system is proposed that is based on random decay in components of the decision process, leading to a steep age dependent increase in the probability of error. The age incidence of those autoimmune diseases which peak in early and middle life can be explained as the resultant of two processes: an exponentially falling curve of incidence of first contact with common bacteria, and a rapidly rising error function. Epidemiological data on the variation of incidence with social class, sibship order, climate and culture can be used to predict the likely site of carriage and mode of spread of the causative bacteria. Furthermore, those autoimmune diseases precipitated by common viral respiratory tract infections might represent reactions to nasopharyngeal bacterial overgrowth, and this theory can be tested using monoclonal antibodies to search the bacterial isolates for cross reacting antigens. If this model is correct then prevention of autoimmune disease by early exposure to low doses of bacteria might be possible.

There are many autoimmune diseases and there are probably many causes of them. Drugs are known to precipitate autoimmune disease by several mechanisms.[1] It is thought that viruses might precipitate autoimmune disease[2 3] by changing the host immune response, releasing sequestered antigens, changing host antigens, cross reacting with host antigens, or by the development of antibodies to host antiviral idiotypes. Bacteria can also promote autoimmune disease. *Streptococcus pyogenes* infection can lead to the development of immune complex mediated glomerulonephritis and rheumatic fever.[4] The former is thought to arise because the bacteria secrete products which form the antigenic component of immune complexes, while the latter is probably a result of cross reaction between bacterial surface antigens and host tissues, particularly heart muscle. The association between streptococcal infection and autoimmune disease was recognised because *S pyogenes* produces local tissue damage as well as the more remote immune mediated damage. In this paper I will argue that immune responses to normal commensal bacteria, which do not normally produce clinical disease, might be responsible for some examples of autoimmune disease. The approach depends critically on statistical decision theory, the general principles of which apply to all decisions made in uncertainty.[5]

Accepted for publication 9 July 1986

Elements of decision theory

Consider an ideal detection system attempting to differentiate a signal (h_1) of constant intensity Y, and assume that in the absence of the signal (h_0) there is background noise, which is normally distributed with mean intensity μ and variance S (fig 1). If, in a defined observation interval, h_1 and h_0 are equally likely then the maximum number of correct responses will be obtained by regarding all intensities above x (fig 1) as h_1 and those below as h_0. If, however, the a priori probability of h_1 exceeds that of h_0 the decision criterion should be reduced below x. Equally, if the rewards associated with correct responses and the costs associated with incorrect responses are not equal then the decision criterion needs to be modified.

According to decision theory,[5] the optimum response (the response that maximises expected value) is to accept h_1 rather than h_0 when

$$L_{10}(e) > B$$

where $$L_{10}(e) = \frac{P[e|h_1]}{P[e|h_0]}$$

and $$B = \left[\frac{V_{00} + V_{01}}{V_{11} + V_{10}}\right]\frac{P[h_0]}{P[h_1]}$$

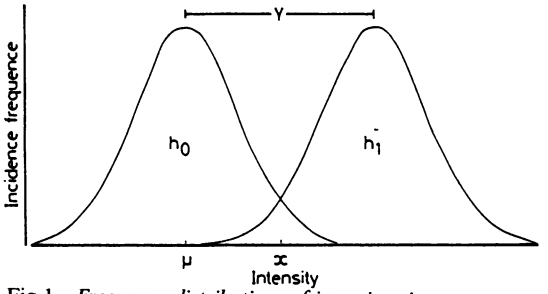

Fig 1 *Frequency distributions of intensity given background noise (h_0), which is normally distributed with mean μ and signal (h_1) of constant intensity Y, which is superimposed on background noise. If h_0 and h_1 are equally likely then a decision criterion placed at x will minimise, but not eliminate, errors.*

L_{10} (e) is the likelihood ratio of event e for h_1 relative to h_0.
P [e|h_1] is the probability of event e given h_1
P [e|h_0] is the probability of event e given h_0
P [h_0] is the a priori probability of h_0
P [h_1] is the a priori probability of h_1
V_{00} is the value of choosing h_0 given h_0
V_{01} is the cost of choosing h_1 given h_0
V_{11} is the value of choosing h_1 given h_1
V_{10} is the cost of choosing h_0 given h_1
H_1 is the decision to accept h_1
H_0 is the decision to accept h_0

There is no single correct answer but only an optimum strategy to maximise expected value. It is clear from fig 1 that any strategy that leads to an increase in the probability of accepting h_1 given h_1 [$P(H_1|h_1)$] will lead to an inevitable increase in the probability of accepting h_1, given h_0[$P(H_1|h_0)$]. The relation between $P(H_1|h_1)$ and $P(H_1|h_0)$ is plotted in fig 2 for two values of Y where $Y_1 > Y_2$. For a constant Y, the curve is generated by varying the value of B (changing the decision criterion).

The analysis so far applies to a simple idealised detection problem. If a decision is to be made, however, between two states h_1 and h_0 then no matter how complex the task and no matter how much information is to be sifted and integrated, the optimal decision rule will depend on likelihood ratio—that is, a simple ratio of the probability of all the evidence given h_1 divided by the probability of obtaining that evidence given h_0. Thus the general principles outlined above apply to all decisions made in uncertainty.

The microbial flora

The body surfaces, particularly the gastrointestinal and upper respiratory tracts, have a complex microbial flora.[6] Although it is often thought that these bacteria are harmless commensals, there is, in fact, clear evidence that many if not most are potentially

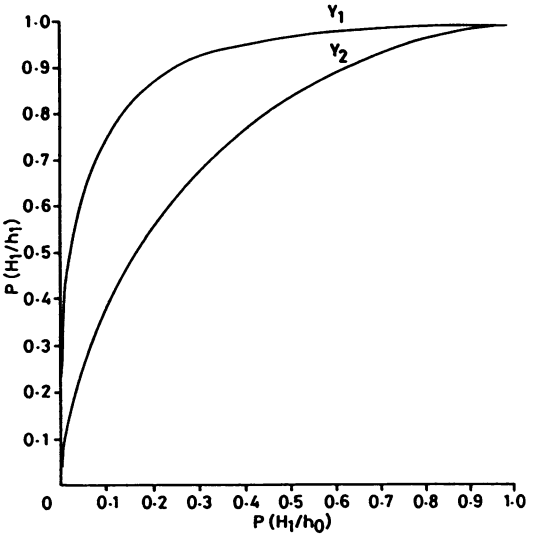

Fig 2 *Probability of true positive $P[H_1/h_1]$ plotted against probability of false positive $P[H_1/h_0]$ for two values of Y where $Y_1 > Y_2$.*

pathogenic. For example, most cases of bronchopneumonia in the elderly are caused by bacteria aspirated from the upper respiratory tract. Perforation of gastrointestinal viscera can lead to peritonitis and septicaemia. Bacterial endocarditis can be caused by organisms such as *Streptococcus viridans* and *Staphylococcus albus*, which are often regarded as harmless commensals. Furthermore, sites of commensal microbial flora can provide temporary or long term residence for a variety of clearly pathogenic bacteria such as *Staph pyogenes, S pyogenes, Haemophilus influenza* type B, encapsulated pneumococci, *Clostridium* sp, *Neisseria meningitidis* etc. This potential for harm can be further extended; bacteria can produce harmful and perhaps lethal toxins (Stephen J, Pietrowski RA, Van Noostrand Reinhold UK, 1981) which might diffuse into the bloodstream. This has been dramatically illustrated with the recent recognition of the toxic shock syndrome caused by toxins produced by certain strains of *Staph pyogenes*.[7 8] This condition can produce severe multisystem disease and sudden death in otherwise healthy adults.[9] It has been known for many years, however, that many staphylococci produce a range of haemolysins which disrupt lipoprotein membranes and are lethal to experimental animals after intravenous injection (Stephen J, Pietrowski RA, Van Noostrand Reinhold UK, 1981). Streptococci,[10] including the ubiquitous viridans group, and clostridia are known toxin producers. Furthermore, bacteria can acquire the capacity to code for toxins on their DNA by plasmid transfer or bacteriophage infection, and this might underlie the toxigenic capacity of *Escherichia coli* (Glass RE. gene

function; Croom Holm Ltd, 1982) corynebacteria, and staphylococci.[11]

A host immunological response to contain bacterial growth on body surfaces and prevent tissue invasion and toxaemia is therefore required. There is a further problem: cross reaction between bacterial surface antigens and host tissues is widespread. Thus *E coli* reacts with antibodies to insulin and brain gangliosides[12]; *Yersinia enterocolitica* have thyrotropin receptors[13 14]; streptococci have surface components cross reacting with HLA antigens[13]; antibodies to chorionic gonadotrophin react with *S epidermidis, E coli,* and *Pseudomonas maltophilia.*[15] Glycopeptides from human and rat brains share antigenic determinants with meningococcus group B and *E coli.*[16] Monoclonal antibodies to nicotinic acetylcholine receptors react with proteins in *E coli, Proteus vulgaris,* and *Klebsiella pneumonia.*[17]

Consequently, the immune system is faced with a complex detection problem, as it needs to classify bacterial antigens into foreign or non-cross reacting (equivalent to group h_1) and self or cross reacting (group h_0) and respond appropriately. No matter how sophisticated the detection system, and no matter how much information is sampled, there must be some residual uncertainty in this distinction. Indeed, as the compatibility of an antigen with an antibody depends on complementary electron cloud shapes and the position of an electron is a random variable, therefore the electron cloud shape is also a random variable, albeit with smaller variance. The response H_1 or H_0 will depend on information (e) gained about the antigen, on the perceived a priori probability of the antigen being self or not self, and on the perceived values V_{00}, V_{11}, V_{10}, V_{01}. Perceived a priori probability will be influenced by previous experience of how frequently cross reacting antigens are found on bacteria, and whether certain types of bacteria or bacteria acting in certain ways are more or less likely to express cross reacting antigens. Perceived values will be influenced by assessment of the likely pathogenicity of the organisms. These calculations represent further decisions made in uncertainty. V_{01} will be a measure of the harm arising by committing a clone of B or T lymphocytes to attack a self antigen, while V_{10} is the penalty of not responding and increasing the chance of bacterial infection. It is possible that once a clone of cells is committed to an antigen (H_1) then the response is permanent, while a failed response (H_0) could be subsequently rectified. It is also possible that there are levels of lymphocyte control which could delete autoaggressive clones, but these control systems will also be subject to uncertainty so there is always a finite change of a permanent autoimmune response following contact with surface bacteria that have antigens in common with the host.

The age incidence of autoimmune disease

Consider the problem of an ideal detection system attempting to distinguish self (h_0) from non-self antigens (h_1) on bacteria. The detection system must in some way obtain samples of information (e) by analysing the antigen along one or many dimensions. It will then be necessary to compute the probability of the totality of information obtained given h_1 [$P(e|h_1)$] and h_0 [$P(e|h_0)$] to determine the likelihood ratio. Finally, the decision made will depend on the value of B, which will be influenced by perceived utilities V_{00}, V_{01}, V_{11}, V_{10} and perceived a priori probabilities. To simplify the model let us assume that a large number [N] of samples of information (e) are obtained from each antigen and transformed on to some linear scale and that the mean of the observations ($\bar{e}$) is determined. Then the probability distribution $\bar{e}|h_1$ and $\bar{e}|h_0$ will tend to have a normal distribution if N is large (central limit theorem). The variance of the distribution will be inversely proportional to N (sampling theorem). The problem is therefore analagous to that in fig 1, with overlapping distributions $\bar{e}|h_1$ and $\bar{e}|h_0$ and a decision criterion (x) whose position is determined by the value of B.

The performance of any complex system must decay with time. This law also applies to biological systems, even though repair systems might limit the rate of decay. A first approximation would be to allow N to decrease at random with time so that the function Nexp $(-kt)$ determines the variance of $\bar{e}|h_0$ and $\bar{e}|h_1$ (fig 1). Then the probability of $H_1|h_0$ [$P(H_1|h_0)$], which is the area under the frequency distribution of $\bar{e}|h_0$ to the right of the decision criterion, will increase with age as the variance of $\bar{e}|h_0$ increases. The time when immune responses to bacteria are set, however, is when the organisms are first encountered. For most common bacteria, this will be early in life, and the age distribution of first contact with any particular organisms will show a roughly exponential curve falling from birth with a half life determined by the prevalence of the bacterium. Thus the age incidence of onset of autoimmune disease arising as a consequence of this contact will be the resultant of a rising curve $P(H_1|h_0)$ interacting with a falling exponential curve. Fig 3 shows a representative result from the model. The exponential curve has a half life of four years, while P $(H_1|h_0)$ is determined by a decision criterion moving from 5·5 standard deviations from the mean at birth at a rate determined by constant k in the function Nexp $(-kt)$. The incidence of the disease in this example rises to a peak at 32 to 36 years, and about 65 cases will arise in a birth cohort of 500 000. If the rate of circulation of the bacteria decreases so that the half life lengthens then the incidence of the disease would rise substantially. Equally, a relatively

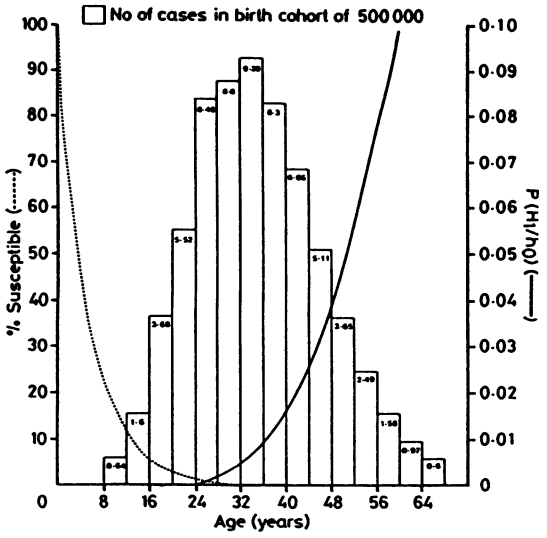

Fig 3 *Theoretical age distribution histogram of autoimmune disease produced by rising error function* $P[H_1/h_0]$ *interacting with falling curve of susceptibility.*

small increase in the rate of bacterial circulation would cause a substantial reduction in disease incidence, and if this model is correct efforts to ensure that children meet the putative bacteria early could in theory virtually eliminate the condition.

This type of age incidence curve with a peak in early or middle life is seen in many putative examples of autoimmune disease,[18] including multiple sclerosis,[19] thyrotoxicosis, insulin dependent diabetes mellitus,[20] ulcerative colitis and Crohn's disease.

Discussion

There is clear evidence to show that many bacteria of the normal microbial flora are capable of causing disease either by tissue invasion or by elaborating toxins. There is also evidence of extensive cross reaction between bacterial antigens and host tissues. The immune system is faced with an exceedingly complex task in determining responses to bacterial antigens so as to avoid both infection and autoimmune disease. The fact that it is successful in most cases indicates that the decision strategy adopted is close to the ideal, so that modelling the decision process using the properties of an ideal detector is a reasonable approximation. If the immune system deviates from an optimal strategy then this will only amplify errors that are inherent in the decision process and which have been identified using the ideal detector model. It is not assumed that there is a structural entity corresponding to the ideal detector, but that the integrated action of the many and diverse arms of the immune system approximates to an optimum strategy. As the immune response to bacteria is decided in uncertainty there will be a finite change of the error $H_1|h_0$, and it can be predicted that at least some examples of autoimmune disease will arise as a response to microbial flora antigens. Obviously this in no way implies that this is the only way in which autoimmune disease arises. It is certainly possible that autoimmune disease could arise as a result of cross reacting antigens present on viruses. Interestingly, while viruses have a coding capacity of between 3 and 50 kilobases of DNA or RNA, bacteria have up to 5000 kilobases of DNA.[21] Given the large number of bacteria resident on the body surface, the antigenic challenge posed by bacterially coded polypeptides is many 100-fold more complex than even a lifetime's accumulation of viral infections.

The age incidence of many examples of autoimmune disease, which show a peak in early or middle life, can be explained using this model. The assumptions are that the immune response is set at first contact and that the performance of the ideal detector decays at random with time. The latter assumption means that a simple exponential decrease in the number of samples of information (N) leads to a steep age dependent increase in $P[H_1|h_0]$. Burch has shown that there is an approximate linear relation between log incidence and log age for many diseases of aging[22-24] including malignant disease, degenerative diseases, and cardiovascular disease. It is therefore of interest that there is also an approximate linear relation between log $P[H_1|h_0]$ and log age over a wide range of values. The concept of random decay in the components of a decision process leading to increasing probability of errors of function might have a wider application.

If the age at first contact with a bacterium determines the probability of autoimmune disease then communities in which the bacteria circulate more rapidly will have a lower incidence of the disease. The spread of enteric bacteria might be influenced by social class, sibship order, and climate. In this respect the observation that multiple sclerosis is more common in higher social classes,[25] first born children,[26] and colder climates[27] is of considerable interest. On the other hand, the spread of respiratory tract bacteria might be influenced by sibship order, but is less likely to be affected by social class, and might actually increase in colder climates. Consequently, the epidemiological features of autoimmune disease could give a clue to the normal mode of spread of the initiating organisms.

There is some evidence to suggest that common viral respiratory tract infections can precipitate autoimmune disease.[28-30] When this is observed it is usually assumed that the virus acts directly, either by

presenting new antigens, changing host antigens, or interfering with lymphocyte function. There are several problems associated with this explanation. Different viruses seem to be capable of initiating the same disease, and waves of viral infection in a community only precipitate disease in a fraction of the population at risk. Furthermore, although the incidence of any one virus fluctuates considerably from year to year, the incidence of autoimmune disease shows only small yearly fluctuation.[30] Viral infections of the respiratory tract also disturb the normal microbial flora and lead to overgrowth of staphylococci, streptococci, or Gram negative bacilli.[31] As the antigenic challenge posed by this bacterial overgrowth is many fold greater than that of the initiating virus, it is not unreasonable to suggest that at least some examples of autoimmune disease might arise as a response to the bacterial antigens.[32] If by chance the first encounter with a particular bacterium follows a viral infection and the bacteria overgrow in the nasopharynx then the host is faced with a potentially threatening infection necessitating a swift response. According to decision theory, swift responses mean that less information is sampled (decreased N), potentially threatening infections increase V_{10}, and both these changes increase $P[H_1|h_0]$. This idea would explain why autoimmune disease only arises as a very rare complication of a common viral infection and why different viruses could precipitate the same disease. Furthermore, as the spread of bacteria that grow on the body surface is less affected by host immunity than is the spread of viruses this would explain why autoimmune disease shows less yearly fluctuation than does epidemic viral disease.

Autoimmune disease may be acute, subacute, or chronic but often shows a characteristic remitting and relapsing course. This is understandable in terms of this hypothesis, as the microbial flora undergoes constant change under a variety of influences, not least of which are viral infections of the host and bacteriophage infections of the bacteria. If the activity of autoimmune disease is related to antigenic challenge then it will wax and wane as the bacterial flora changes.

One aspect of autoimmune disease that has received a great deal of attention is the role of genetic constitution, in particular HLA phenotypes.[33] Ninety five per cent of whites with insulin dependent diabetes mellitus express the class II HLA alleles DR3 or DR4, or both.[34] These alleles are only expressed in 40% of the general white population. Similar associations between HLA phenotype and other examples of autoimmune disease have been reported.[33] Although genetic factors are important, however, concordance in identical twins for diseases such as insulin dependent diabetes mellitus[29 30] is less than 100%, indicating the importance of environmental factors. The HLA complex is a multigene family on the short arm of chromosome 6 concerned with self- and non-self discrimination. Class I HLA molecules participate in target cell recognition by suppressor and cytotoxic T cells, while class II HLA molecules, which are expressed on the surface of antigen presenting cells, participate in antigen recognition by T helper cells. The relation between HLA antigens, autoimmunity, and infection is undoubtedly complex. If some examples of autoimmune disease arise in response to bacterial antigens as suggested in this paper then as HLA molecules have a central role in T cell recognition of foreign antigens, it is not surprising that HLA phenotypes influence the risk of autoimmune disease. Other ways in which this association might arise include increased susceptibility to certain viral infections and defective T cell suppressor activity, allowing the proliferation of autoreactive clones. Genetic factors could influence the decision theory model in several ways. As genetic factors influence the structure of cell surface antigens they will determine the bacterial antigens to which the individual is susceptible. In fact, the composition of the bacterial flora is influenced by genetic factors as the colonisation of epithelial surfaces by bacteria depends on specific bacterial adhesions combining with cell surface molecules.[6] Furthermore, the incidence of autoimmune disease will vary between subjects who operate different risk strategies. Different perceptions of values [V_{00}, V_{10}, V_{01}, V_{11}] and a priori probability will influence B, and thereby vary the probability of error $[H_1|h_0]$. Indeed, the higher incidence of autoimmune disease in women might be explained in this way as men seem to be more prone to infection and malignant disease, which can be regarded as the consequence of the converse error $(H_0|h_1)$.

The hypothesis presented is amenable to laboratory investigation, although testing it will be a formidable task. Epidemiological data can be used to determine the likely mode of spread of the causative bacteria. For those diseases that are precipitated by common viral respiratory tract infections the bacterial flora in the upper respiratory tract should be determined to see of there is any constant pattern preceding onset or exacerbations of the disease. If there is, monoclonal antibodies can then be used to search the bacterial isolates for antigens that cross react with host tissues.

References

1 Russell AS. Drug induced autoimmune disease, in Autoimmunity edited by EJ Holborow. *Clinics in Immunology and Allergy* 1981;1:57–76.

2 Mims CA. Immunopathology in virus disease. *Philosophical Transactions of the Royal Society of London, series B* 1983;303:189–98.

3 Schoenfield Y, Schwartz RS. Immunologic and genetic factors in autoimmune disease. *New Engl J Med* 1984;**311**:1019–29.
4 Zabriskie JB. Immunopathological mechanisms in bacterial-host interaction. *Philosophical Transactions of the Royal Society of London, series B* 1983;**303**:177–87.
5 Green DM, Swets JA. *Signal detection theory and psychophysics*. London: John Wiley and Sons, 1966.
6 Mackowiak PA. The normal microbial flora. *New Engl J Med* 1982;**307**:83–93.
7 Todd J, Fishaut M, Kapral F, Welch T. Toxic shock syndrome with phage group 1 Staphylococci. *Lancet* 1978;ii:1116–8.
8 Schrock CG. Disease alert. *JAMA* 1980;**243**:1231.
9 Morris JA. Tampon associated staphylococcal infection and sudden death. *Lancet* 1983;i:772.
10 Wannamaker LW. Streptococcal toxins. *Rev Infect Dis* 1983;**5(suppl 4)**:723–32.
11 Anonymous. Toxic shock syndrome and lysogeny. [Editorial]. *Lancet* 1983;**1**:1420–1.
12 Le Roith D, Shiloach J, Roth J, Lesniak MA. Insulin or a closely related molecule is native to Escherichia coli. *J Biol Chem* 1981;**256**:6533–6.
13 Anonymous. Thyroid disease and antibodies to Yersinia. [Editorial]. *Lancet* 1977;ii:734–5.
14 Weiss M, Ingbar SH, Winblad S, Kasper DL. Demonstration of a saturable binding site for thyrotropin in Yersinia enterocolitica. *Science* 1983;**219**:1331–3.
15 Marcio T, Cohen H, Segal SJ, Koide SS. Production of choriogonadotrophin like factor by micro-organisms. *Proc Natl Acad Sci USA* 1979;**76**:6622–6.
16 Finne J, Leionon M, Makela PH. Antigenic similarities between brain components and bacteria causing meningitis: implications for vaccine development ande pathogenesis. *Lancet* 1983;ii:355–7.
17 Stefansson K, Dieperink ME, Richman DP, Gomez CM, Marton LS. Sharing of antigenic determinants between the nicotinic acetylcholine receptor and proteins in Escherichia coli, Proteus vulgaris, and klebsiella pneumoniae. Possible role in the pathogenesis of myasthenia gravis. *New Engl J Med* 1985;**312**:221–5.
18 Beeson PB, McDermott W, Wyngaarden JB. *Cecil textbook of medicine*. London: WB Saunders Company, 1979.
19 Kinnunen E. Multiple sclerosis in Finland: evidence of increasing frequency and uneven geographic distribution. *Neurology* 1984;**34**:457–61.
20 Bloom A, Hayes TM, Gamble DR. Register of newly diagnosed diabetic children. *Br Med J* 1975;iii:580–3.
21 Murray K. New routes to drugs, diagnostic agents and vaccines. *Lancet* 1984;ii:1194–8.
22 Burch PRJ. Mutation, autoimmunity and ageing. *Lancet* 1963;ii:299–300.
23 Burch PRJ. Cardiovascular diseases: new aetiological consideration. *Am Heart J* 1964;**67**:139–340.
24 Burch PRJ. *The biology of cancer: a new approach*. Lancaster: MTP Press, 1976.
25 Posksanzer DC, Schapira K, Miller H. Multiple sclerosis and poliomyelitis. *Lancet* 1963;ii:917–21.
26 James WH. Multiple sclerosis and birth order. *J Epidemiol Community Health* 1984;**38**:21–2.
27 Kurland LT. The epidemiologic characteristics of multiple sclerosis. In: Vinken PJ, Bruyn GW, eds. *Handbook of clinical neurology. Multiple sclerosis and other demyelinating diseases*. Vol 9. Amsterdam: North Holland, 1970.
28 Sibley WA, Bamford CR, Clark K. Clinical viral infections and multiple sclerosis. *Lancet* 1985;i:1313–5.
29 Barrett-Connor E. Is insulin dependent diabetes mellitus caused by Coxsackievirus B infection? A review of the epidemiologic evidence. *Rev Infect Dis* 1985;**7**(2):207–15.
30 Gamble DR. The epidemiology of insulin dependent diabetes with particular reference to the relationship of virus infection to its aetiology. *Epidemiological reviews*. 1980;**2**:49–69.
31 Ramfrez-Ronda CH, Fuxench-Lopez Z, Nevarez M. Increased pharyngeal bacterial colonisation during viral illness. *Arch Intern Med* 1981;**141**:1599–603.
32 Morris JA. Clinical viral infections and multiple sclerosis. *Lancet* 1985;ii:273.
33 Pollack MS, Rich RR. The HLA complex and the pathogenesis of infectious diseases. *J Infect Dis* 1985;**151**:1–8.
34 Eisenbarth GS. Type I diabetes mellitus: a chronic autoimmune disease. *New Engl J Med* 1986;**314**:1360–7.

Requests for reprints to: Dr JA Morris, Consultant Pathologist, Department of Pathology, Lancaster Moor Hospital, Lancaster LA1 3JR, England.